Carnegie-Rochester Conference Series on Public Policy 46 (1997) 1-43
North-Holland

Human capital development and welfare participation*

Robert A. Miller†
Carnegie Mellon University, GSIA, Pgh., PA 15213
and
Seth G. Sanders
Carnegie Mellon University, Heinz School, Pgh., PA 15213

Abstract

This paper uses a sample from the National Longitudinal Survey of Youth (NLSY) to conduct an empirical investigation of how the Aid to Dependent Families with Children Transfer Program (AFDC) affects two types of human capital formation, namely, the educational attainment of young women, and their subsequent performance in the labor market. The analysis begins by confirming two stylized facts about AFDC, namely, the high rates of nonparticipation in AFDC by eligible females, and the negative duration dependence of those who enroll in AFDC. Then the paper uses reduced-form econometric techniques to demonstrate that educational attainment is not significantly affected by the level of welfare support provided by different states, but that hours worked in the labor force declines. The latter parts of the paper estimate a dynamic model of discrete choice which incorporates the features of wage growth through experience on the job, and preferences which adapt to labor supply and welfare experience, to explain nonparticipation in AFDC among eligible families and the observed negative duration dependence of families who enroll. Both model features are found to play important reinforcing roles in explaining these two empirical regularities.

**Correspondence to:* Robert A. Miller, Carnegie Mellon University, GSIA, Pittsburgh, PA 15213.

†Chandra Das, Terra McKinnish, and Rafael Mendible provided indispensable research assistance. We also thank our discussant, Ariel Pakes, the editor, Bennett McCallum, and Vesna Prasnikar for their helpful comments.

PII: S0167-2231(97)00002-X

1 Introduction and summary

The Aid to Dependent Families with Children Transfer Program (AFDC) has existed for more than half a century, growing in both generosity and population coverage in the two decades immediately following World War II. Although the real benefits to participants vary with state legislation, the generosity of the program peaked about twenty years ago, declining in generosity until a decade ago and has remained stable in real dollar terms since then. By the late 1980s benefits to AFDC participants had declined from their peak to around the same real dollar value as was available to participants in 1960. When added to the benefits available from related programs for which AFDC participants qualify, Medicaid (introduced in 1965 to subsidize medical health care to the poor) and the Food Stamp program, the gross benefits to families on Public Assistance have increased quite substantially since the 1970s. Although AFDC is hardly a new program, only with the availability of cross-sectional and panel data sets in the last 25 years has it attracted much attention from social scientists seeking to measure its quantitative impact.[1]

Two empirical regularities are found across many studies of participation in the AFDC program. First, many women who would seemingly benefit from public aid do not participate in welfare programs. Studies of AFDC participation estimate that between 20 and 40 percent of eligible households fail to participate in AFDC.[2] Many other female heads of families, through a great deal of work effort, earn more than AFCD allows for participation and therefore do not use AFCD. Some have argued that these women should also be counted as nonparticipants, because if these women valued leisure even modestly they would be better off (in a utility sense) on welfare. While the definition and the associated estimates of nonparticipation vary across studies, all studies conclude there are many women who would benefit from receiving AFDC but who choose not to participate in the program. A second regularity found across studies is that exit rates from AFDC fall with time spent on AFDC, that is, spells of AFDC exhibit negative duration dependence.[3] O'Neill, Wolf, Bassi, and Hannan (1984) find this in three sep-

[1]See Moffitt (1992) for a recent survey of the incentive effects of the welfare system.

[2]Moffitt (1983) using data from the CPS estimates that the AFDC participation rate among female-headed families was 63% in the early 1980s. Ruggles and Michel (1987), using repeated CPS panels and a microsimulation model to impute household assets (not directly observable on the CPS) estimate the annual participation rate for AFDC in the mid-1980s was 75% to 80% among eligible women.

[3]Negative duration dependence refers to the fact that time in an AFDC spell lowers the hazard rate of exit from that spell. Thus, negative duration dependence implies that the probability of remaining on welfare for another period increases with time spent on welfare.

arate data sources.[4] Blank (1989) finds that at least for part of the AFDC recipient population, the observed negative duration dependence in AFDC spells is not simply a statistical artifact of unobserved heterogeneity among recipients.

Much of this empirical work has been organized around a static framework where entry into welfare or exit from welfare is a function of individual characteristics at a point in time. Usually at a point in time individual characteristics are thought to be exogenous for the purpose of present decisions. If there were no long-term consequences, this simplification would provide a convenient framework in which to couch empirical findings. In this case the effects of any proposed change to the system could be neatly partitioned into two pieces, the immediate effects on eligibility, and the effects on the take-up rate among those who are eligible. For example, in studying the effects of AFDC on labor supply, researchers have confirmed a broader finding in the labor-supply literature, that the labor-supply elasticity is low, AFDC recipients hardly responding to the benefit reduction ratio (set by the national government), and that most of the variation in labor supply and labor-force participation observed over the years among those receiving AFDC benefits is due to changes in the eligibility rules.

The problems associated with treating each period independently of the other periods are well-known. When preferences are not additively separable over time and/or markets are incomplete, current decisions have future ramifications, and consequently changing a welfare scheme, even temporarily, has long-term effects. There are three types of human accumulation that immediately come to mind when contemplating the provisions of AFDC. Since family income helps determine eligibility requirements, one could investigate the effects of AFDC on the adult composition of families, particularly marriage and divorce. Second, the costs of bearing and raising children are, by design, reduced by AFDC (Medicaid and Food Stamps), and this may affect sexual behavior, contraceptive choices, and fertility. Third, AFDC might reduce the incentives to invest in schooling and work experience, both of which raise productivity in the labor market. Because none of these issues can be neatly parceled into their one-period effects, a dynamic framework is required to analyze them, but it is challenging to conduct an empirical analysis which produces findings that are robust to alternative assumptions about the means for transferring resources between periods and how much preferences are inter-temporarily linked.

[4] O'Neill et al. confirm this in data from the Panel Study of Income Dynamics, data from the National Longitudinal Survey of Labor Market Experience, and in AFDC caseload records. Blank (1989), using data from the Seattle/Denver Income Maintenance Experiment (SIME/DIME), replicates the finding. This pattern is also present for the women in the National Longitudinal Survey of Youth (NLSY), the principal data used in this study.

In contrast to much of the previous empirical work aluded to above, focusing on the dynamic issues such as human capital accumulation for the labor market does not lend itself to dividing the effects of AFDC between the eligibility and the take-up rate in any given year. The reason is that conscious decisions about educational attainment made in teenage years before a person becomes eligible may indirectly affect eligibility later on. Therefore, in any one year, those who are ineligible for welfare may take actions that either increase or reduce the likelihood of becoming eligible and receiving benefits in future years.

This paper has two aims. First, it is an empirical investigation of the effects of AFDC on female schooling and subsequent work experience, including labor supply and wages. Second, the paper addresses whether these human capital mechanisms can explain the patterns of nonparticipation and duration dependence found across many studies. Section 2 describes the National Longitudinal Survey of Youth (NLSY) data used in our empirical analysis. There we also document the fact that the NLSY displays the same empirical regularities found elsewhere: a substantial fraction of eligible women apparently do not receive AFC and the probability of leaving AFDC declines with time spent on the program (called negative duration dependence). Section 3 undertakes an exploratory data analysis to investigate ways in which taking welfare assistance may retard investment in human capital. In particular we focus on how the allocation of time between formal education, labor supply, and leisure might be affected by increasing the generosity of AFCD programs. We conclude from the exploratory data analysis that while achievement in formal education by young women is insensitive to changes in the benefits available from AFDC, future labor supply appears to decline with the generosity of the program available to a woman as a teenager. Furthermore, our findings suggest that the reason for this decline is not that the program itself has a direct effect, but that it discourages women from early labor-force experience and that labor supply itself is serially correlated.

Section 4 presents a dynamic structural model of discrete choice which incorporates this human capital mechanism in a model that allows for welfare stigma and habit formation in work or welfare to explain the observed patterns of welfare participation. The static stigma model is nested within the more general dynamic model, formulated so that it is possible to reject both forward-looking behavior, stigma, and/or tolerance as mechanisms of welfare nonparticipation and time dependence.

The results of the structural estimation are reported in the last three sections of the paper. In Section 5 we estimate a wage equation and find that past labor-force participation adds to marginal productivity; the more recent the work experience, the larger its impact on wages. This finding establishes the importance of dynamics, since forward-looking agents would make their labor-supply decisions recognizing that the current wage captures only part of the financial benefit from working. Then in Section 6 we undertake non-parametric estimation with the aim of illustrating the patterns in work-force and welfare participation, conditional on having optimally chosen one of the four work-welfare combinations that arise from our structural model. Our main result is that negative duration dependence in welfare participation overwhelms the human capital element from labor supply (reported in Section 5). Once a person accepts welfare, the probability that she will choose not to work and will continue with AFDC in future periods hardly depends on whether she was initially working or not.

This naturally raises the question whether the human capital element in wages is offset by a distaste for work that comes with labor-supply experience. In fact we find the opposite result in Section 7, where our estimates of the structural parameters characterizing preferences over consumption and leisure choices are reported. Past and current labor are found to be complements, suggesting that even in the absence of age- and experience-driven wage growth, women develop a taste for work, or in the language of household economics, acquire skills that help them manage production in the household and labor sectors simultaneously. In addition, we find that stigma against welfare factors into preferences. The qualitative nature of these estimates is not sensitive to the subjective discount factor, which is hard to estimate precisely. Having admitted that, however, we remark that our point estimates of this parameter suggest that young women are forward-looking people who take dynamic considerations into account as they plan their futures.

2 Data for the study

2.1 *Data sources*

The principal data source for this study is the National Longitudinal Study of Youth (NLSY). The NLSY follows 6283 young women, ages 14 to 21 in 1979, from 1978 to 1985. The NLSY oversamples the poor, contains a rich set of information on family background, has monthly information on labor force and welfare participation, and has information on the exact geographic location of respondents. Knowing the county and state of residence of survey respondents allows us to exploit several additional data sources to more accurately describe the welfare and work constraints that respondents face. Estimates of the welfare payments per child in the household and the AFDC tax rate on earnings were used to characterize the generosity of each state's welfare program.[5] In addition a measure of sectoral employment growth was constructed for the county of residence of each woman in the NLSY from es-

[5]See Fraker, Moffitt, and Wolf (1985).

timates of annual employment by sector derived from administrative records of each state's unemployment insurance program. It is used in the analysis below to reflect demand for the woman's labor.[6]

The latter parts of the paper focus on the use of welfare by the 2073 women in the NLSY who were age 13 to 16 on January 1, 1978. Information on the older women is used, but only to construct forecasts of what the younger women might expect their future to hold in response to the choices that they currently make. Concentrating on the welfare and work behavior of the young has several advantages. While pre-survey information on work, marriage, pregnancies, and births is available, pre-survey information on welfare participation is unavailable. A primary goal of this work is to explain the relationship between human capital accumulation and welfare participation when past welfare participation is also allowed to influence current welfare participation. While pre-survey welfare use is unreported, women first observed at ages 13 to 16 are unlikely to have participated in welfare before the survey stated. In fact, only three women in the sub-sample had welfare spells in progress at the survey's start. Because the spells of young women tend to start with the birth of a child and last a long time, it is also unlikely that these young women have experienced any previous welfare spells. A second advantage of studying the youngest women is that what happens in their teenage years may permanently affect their risk of becoming eligible for the AFDC program. Indeed there is concern about the rate of teenage pregnancy, high school dropouts, and subsequent welfare usage. The sampling plan of the NLSY implies that when only women aged 13 to 16 in 1978 are selected, we still observe 385 women who experience 537 welfare spells during the first eight years of the survey, an ample number to conduct a statistical analysis.

2.2 *Non-participation and duration dependence*

The NLSY data, like many other data sources, show that many women do not participate in AFDC. Table 1 describes the proportion of all women in the NLSY who did not participate in AFDC although eligible, by family status, to do so.[7] In 65.4 percent of the months that women met their state's family status eligibility criteria they did not participate in welfare. Since women must meet both the family status criteria and an income criteria to qualify for the AFDC program, some women who meet their state's family status criteria may have been disqualified from participation by earning more than

[6] This information is collected and compiled for public use by the Bureau of Economic Analysis.

[7] Women who were either unwed mothers or pregnant and living in a state that allowed pregnant women to participate were considered eligible by their family status.

their state allowed for participation.[8] Table 1 shows that in 69.5 percent of the months for which women were eligible for welfare by their family status were also eligible based on their level of income. Therefore, women who met both their state's family status eligibility criteria and their state's income criteria did not participate in AFDC 45.5 percent of the time. These women would have had larger incomes had they participated in the AFDC program.

The table implies 31.2 percent of income eligible nonrecipients worked and earned less than they would have received from welfare. The remaining family eligible nonrecipients could have qualified for welfare by reducing their labor supply. In only 11.6 percent of the months did women who qualified for welfare by their family status (an unwed mother) earn more than double the welfare earnings' cutoff in their state. Given the cost of going to work (such as the loss of leisure, alternative arrangements for child care, transportation, and extra clothing) and the noncash benefits of the welfare program, it seems that most who were eligible by their family status but who did not participate in AFDC would have been better off accepting welfare.

The fact that a large number of eligible women neither participate in AFDC nor work is evidence against the hypothesis that acquiring human capital is the only reason for not becoming a welfare recipient. The fraction of months that this occurs is a function of age. Across all ages, women who were eligible but did not take welfare also did not work 47.8 percent of the time. However, among those older than 26, eligible nonrecipients worked in all but 28.4 percent of the months. In 20.4 percent of the months for this older age group, women worked for less than what their earnings would have been on welfare, and in an additional 26.6 percent of the months these women worked for wages that were greater than the welfare earnings' cutoff by a factor of less than two. As they grow older, women seem to become more inclined towards work rather than accept welfare.

One explanation why a woman neither works nor receives welfare is that she may be ineligible for welfare on other grounds apart from family status and income. Many states have additional requirements for eligibility that take into account the earnings and assets of other family members. The youngest women, who are most likely to live with their parents, are most affected by such state requirements, compared to say 26-year-old women, who are more likely to have moved out of their parental home.

The NLSY also exhibits the typical pattern of duration dependence in welfare spells. Table 2 provides information on the distribution of welfare spells for the 2073 women ages 13 to 16 in 1978. Of the 537 spells on welfare, 323 spells end during the survey while 214 are right censored. Consistent with

[8]The earnings cut-off for each woman was calculated as Cut Off = (dollars per child paid by the State) × (Number of Children of the Woman)/(AFDC tax rate). This is an imputed value of the maximum she could earn and still qualify for AFDC.

Table 1:
Welfare Nonparticipation by Eligible Women and Their Earnings

	All Ages		Greater Than 26 years old	
NOT WORKING	47.8		28.4	
0< 25%	3.7	21.7%	3.3	20.4%
25 - 50%	5.0		3.6	
50 - 75%	6.2		7.7	
75 - 100%	6.8		5.8	
INCOME CUT OFF				
100 - 125%	6.1	18.9%	8.4	26.6%
125 - 150%	5.1		5.7	
150 - 175%	4.2		6.1	
175 - 200%	3.5		6.4	
> 200%	11.6		24.7	
ELIGIBLE BY FAMILY STATUS	65.4%			
ELIGIBLE BY FAMILY STATUS AND INCOME	45.5%			
PROPORTION OF INCOME ELIGIBLE THAT WORKED	31.2%			

findings from other studies, the majority of welfare spells last a relatively short time.[9] For example the cumulative frequency in Table 2 shows that 61.1 percent end within 2 years. However, the rate at which women exit from welfare slows with time spent on the program. Figure 1 plots the hazard rate of exit from welfare at the midpoint of 12-month intervals. This figure shows that while 3.5 percent of the women who have been on welfare for 6 months are estimated to leave at that time, only 1 percent of the women who have been on welfare for 66 months are estimated to leave then. Figure 1 confirms for the NLSY data what is found in many other studies: broadly speaking, the longer women stay on welfare, the less likely they are to leave the program.

Table 2:
The Distribution of Welfare Spells

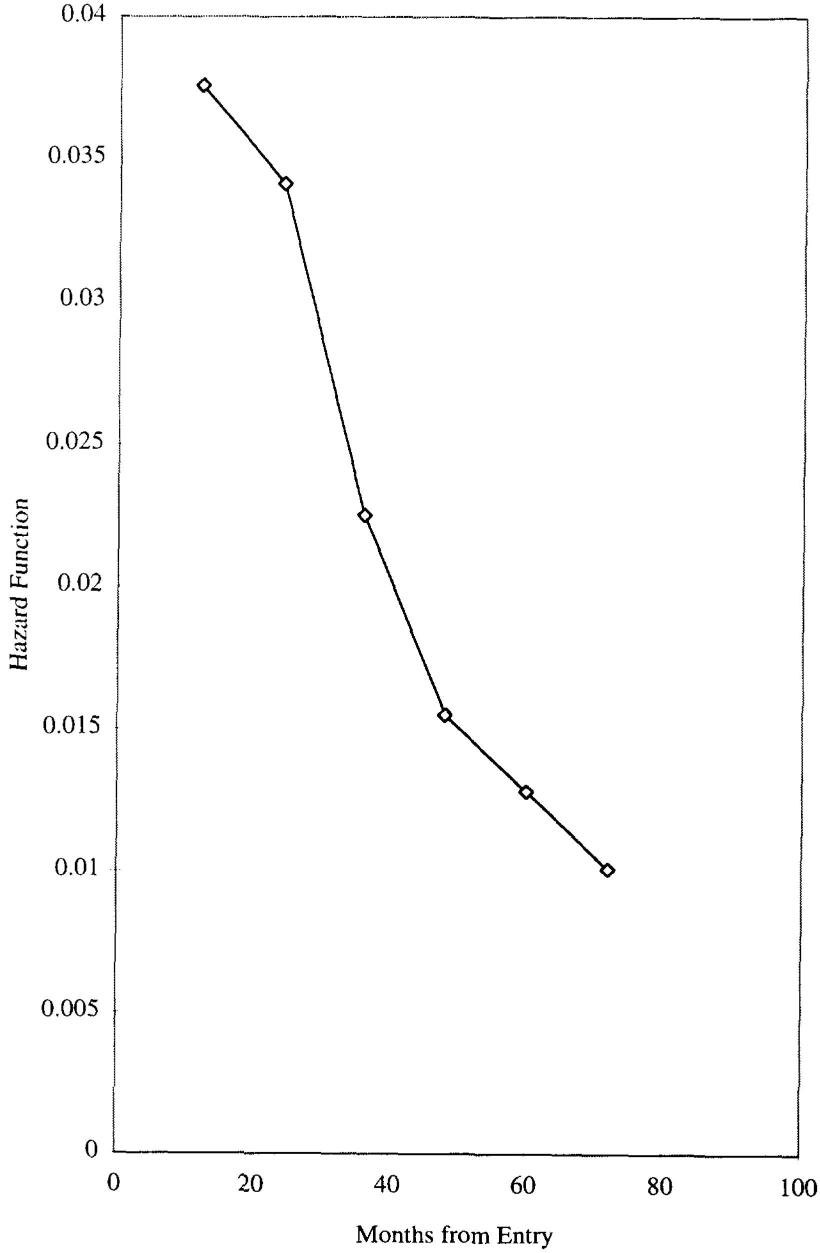
Time	CDF	Number Completed in	Number Censored After
2 MONTHS	9.5	48	47
4 MONTHS	15.5	102	55
6 MONTHS	24.2	133	60
12 MONTHS	47.2	223	97
24 MONTHS	61.1	278	134
36 MONTHS	71.0	308	166
48 MONTHS	76.0	317	183
60 MONTHS	79.2	321	198
72 MONTHS	81.7	323	206
84 MONTHS	81.7	323	211
96 MONTHS	81.7	323	214
TOTAL SPELLS	537	323 (60%)	214 (40%)

Individual differences that persist through time, when not properly accounted for, induce negative hazard even in the absence of duration dependence at the individual level (because those who have been on welfare a long time are disproportionately represented by those least likely to exit welfare). Thus, while decay in human capital may produce dependence on welfare, if experiencing welfare truly changes a woman's market opportunities, it is important to purge the data of the effects of heterogeneity. Before investigating a casual relationship between human capital and welfare participation,

[9]See Blank (1989), for example.

FIGURE 1

Non-Parametric Hazard Estimates of Exit from Welfare



one would like to know if the duration patterns observed in Figure 1 simply reflect differences among women. To answer this question, hazard models were estimated to control for some of the differences among women. The results reported here are for a hazard which is linear in both duration dependence and the controls added to account for heterogeneity. They include two measures of welfare payments in a woman's state of residence (AFDC GUARANTEE and AFDC BREAK EVEN, which is the maximum income qualifying for welfare), age, indicators for race (BLACK and HISPANIC), the mother's education by years of schooling (MOMED), and information about the respondent's childhood family background (INTACT = 1 if she lived with both parents at age 14 and 0 otherwise). Measures of local labor-market growth were then added to the basic model.[10]

Table 3 displays many of the results usually found. Once controlling for the break-even level, higher welfare payments and/or being black lower the likelihood of leaving the program. The significance of the guarantee effect is statistically stronger in this study than in previous studies.[11] Having a mother with higher education, coming from an intact family, or growing older all have little impact on moving women off welfare. The results of including the local employment growth measures are presented in Columns 2 and 3 of Table 3. There also seems to be an independent effect from growth in the service sector. Indeed, while overall growth in employment only weakly raises the changes of getting off welfare, a growing service sector has a strong impact on raising a recipient's probability of leaving welfare. However, just as with the basic set of controls used above, controlling for growth in job opportunities has little impact on the pattern of baseline duration dependence; the coefficient on duration remains negative and significant.

3 Welfare and investment in human capital

This section investigates whether AFDC benefits are negatively correlated with two forms of human capital accumulation, schooling, and experience

[10]In addition, a dummy variable is included in the month that the information from each annual survey ends to control for an irregularity in the data construction. While respondents are consistent in answering questions within a survey, there are often large inconsistencies when information from multiple surveys is joined to construct life histories of respondents. Therefore, welfare spells are more likely to start and stop in months reported in adjacent survey years than in adjacent months reported within the same survey year. This is often labeled the "seam" problem.

[11]Previous studies use the maximum benefit level a state will pay as a measure of generosity of a state's welfare program. Fraker, Moffitt, and Wolf (1985) estimate the effective AFDC tax on earning and the effective payments to recipients directly from data that contains actual payments. These estimates account for differences in state administrative practices that affect the actual levels of payments to recipients. The Fraker-Moffitt-Wolf estimates are used in the models above.

Table 3:
Estimated Hazard Rate Out of the First Spell on Welfare

Variable	**Parameter Estimates (Standard Errors in Parenthesis)**		
INTERCEPT	-0.155 (1.139)	0.409 (1.605)	0.461 (1.620)
DURATION TIME	-2.642 (0.615)	-2.727 (0.619)	-2.711 (0.627)
AFDC BREAK EVEN	0.017 (0.052)	0.018 (0.052)	0.024 (0.052)
AFDC GUARANTEE	-0.361 (0.158)	-0.369 (0.158)	-0.383 (0.160)
AGE	-0.016 (0.036)	-0.622 (0.037)	-0.591 (0.037)
GLACK	-0.651 (0.148)	-0.622 (0.151)	-0.591 (0.156)
HISPANIC	-0.158 (0.180)	-0.155 (0.181)	-0.142 (0.184)
MOMED	0.169 (0.122)	0.170 (0.124)	0.179 (0.125)
INTACT	0.012 (0.129)	0.024 (0.130)	0.019 (0.134)
% CHANGE LOCAL EMPLOYMENT	-	0.311 (0.187)	-0.114 (0.329)
% CHANGE MANUF EMPLOYMENT	-	-	0.080 (0.120)
% CHANGE SERVICE EMPLOYMENT	-	-	0.437 (0.218)
SPIKE	1.946 (0.072)	1.940 (0.073)	1.937 (0.074)

on the job. Specifically, is living in a state at age 17 with generous welfare payments associated with a lower rate of high school completion or fewer hours worked at age 27? Although this part of our study does not have a structural interpretation, it is not confined to the population of women who are eligible for welfare at age 17 (teenage mothers), because AFDC could affect the human capital choices of women who never became eligible for the program under its existing provisions but might have chosen to participate had the program been more generous.

The basic strategy for identifying the effect of AFDC payments on human capital development is to use variation in AFDC payments across states in a regression on measures of human capital development. While most of the literature on the AFDC use this strategy, Moffitt (1994), Jackson and Klerman (1994), and Hoynes (1995) take issue with the implicit assumption that a state government's AFDC generosity is uncorrelated with other unobserved factors that also influence the dependent variable of interest. One plausible mechanism for this correlation is that state legislatures set funding levels for both the state education system and the state welfare system. Since funding is influenced by state wealth, wealthier states may have better school systems and more generous welfare programs. If differences in wealth across states are not fully captured by the statistics available to the researcher (such as per capita), a state's AFDC generosity may be positively correlated with unobserved factors that influence educational attainment. An alternative mechanism for this correlation is that ideology varies across states; more liberal states presumably provide more support to both education and welfare.[12]

Typically, researchers include state fixed effects in regression models to control for unobserved differences across states, which requires several cross-sections or panel data. This approach identifies the effect of AFDC generosity through changes in AFDC generosity over time within each state.[13] Using several cross-sections of the Current Population Survey, Moffitt (1994) investigates the effects of AFDC generosity on female headship, and finds that adding state fixed effects reverses the effect of AFDC generosity for white women from positive and significant to negative and significant. Using data from vital statistics of the United States to investigate the effects of AFDC generosity on fertility, Jackson and Klerman (1994) find that including state

[12]Bane and Ellwood (1985) discuss this source of bias. It can stem from differences across states in social and cultural norms or religious influences.

[13]See Moffitt (1994), Jackson and Klerman (1994), and Hoynes (1995). Note that if the source of endogeneity is through the legislative process, then including state fixed effects can exacerbate the endogeneity issue. If states' funding for both education and welfare is responsive to changes in tax revenues induced by economic conditions, then unobserved factors influencing human capital development may be more highly correlated with changes in AFDC generosity than the level of AFDC generosity.

fixed effects changes the effect on fertility from negative to positive for white women. These findings suggest that the estimated effects of AFDC on human capital accumulation might also be sensitive to whether identification is achieved through variation over time within states, variation across states at a point in time, or both.

Hoynes (1995) also investigates the effects of AFDC generosity on female headship. Using data from the Panel Study of Income Dynamics (approximately 3800 observations annually), she replicates Moffitt's (1994) positive and significant effect of AFDC generosity on female headship in the absence of state fixed effects. However, controlling for state fixed effects, the effects of AFDC generosity become statistically insignificant. While the coefficient drops towards zero, the standard error on the coefficient doubles as well. While including fixed effects purges time invariant differences across states, one also must rely only on differences in the changes in AFDC generosity across states for identification. Over the two decades beginning in 1970 or thereabouts, most states changed real AFDC benefits by leaving nominal benefits unchanged. Since inflation is highly correlated across states, there is therefore limited independent variation in the differences of changes in AFDC generosity across states.

Ideally, we would like to exploit variation in the part of AFDC payments across states that is uncorrelated with unobserved factors influencing the dependent variable, simultaneously purging any part of AFDC payments that is correlated with the unobserved factors. In practice, we estimate models where, after conditioning on the high school completion rate for men in the state, the level of AFDC generosity serves as our source of identification for the effect of AFDC on high school completion of women. Including the high school completion rate of men in the state as an exogenous variable purges all the unobserved factors affecting a women's rate of high school completion that are common to both men and women in the state. In particular, since men and women attend the same schools, legislative decisions within the state over educational funding issues are likely common to both men and women.

Table 4 reports the mean characteristics of men and women in our sample from the NLSY weighted to represent the U.S. population in 1978. The variables AGE 14 through AGE 21 are dummy variables which equal one if the respondent was in the age group and zero otherwise (for example, AGE 14 = 1 if the respondent was 14 years old in 1978, AGE 14 = 0 otherwise). In the regressions below, these serve as controls for differences across cohorts in high school completion rates and hours worked. Two race categories are included, as are two immigration categories (RIMG = 1 if the respondent is an immigrant and PARIMG = 1 if the respondent's parents are both immigrants), and two measures of the family structure of the respondent at age 14 are included (INTACT and indicator variable FEMHEAD = 1 if the respondent lived with her mother only). Several measures of the respondent's parents are included to capture the intergenerational transmission of human capital: the level of educational attainment of the respondent's mother (MOMED) and father (DADED), and whether or not her mother was on welfare in 1978 (FAMWELF78). An ability measure of the respondent, her AFQT score, is also reported, where AFQT is measured in terms of standard deviations from the mean score for the age of the respondent when the exam was administered. Finally, the real AFDC payment (1995 dollars) that a women could have received at age 17 if she became eligible is also reported (RAFDC).[14] The mean guarantee was approximately \$430 per month but there was a large degree of variation across states and time.

Table 5 reports the parameters and standard errors obtained from estimating a logit model of high school completion by women.[15] The first column reports the analysis without state fixed effects, the next column reports the analysis with state fixed effects, and the last one reports the analysis conditioning on the high school completion rate for men in the state. While most covariates have the expected signs, there is very little evidence of an effect of AFDC generosity on high school completion rates. controlling for endogeneity, with either state fixed effects or with men's high school completion rate, we cannot reject the hypothesis that AFDC generosity does not affect high school completion rates. Table 6 reports the findings from running the same regressions on women whose parents' household income was less than \$10,000 a year in 1978 (approximately half the sample). Among the daughters of poor families, where one have might expected the most impact, we cannot detect an effect of AFDC generosity on high school completion rates.

Linear regression models were estimated to gauge the effects of AFDC generosity at age 17 on hours worked by women 10 years later. To control for demands that children place on their mother's time, we included four regressors, denoted KID 1 through KID 4, on the number of children in several age groups (less than 2 years, between 2 and 6 years, 6 to 11, and older than 11). From Table 7 we see that all three estimators (cross-sectional, fixed-effect, and opposite sex) suggest that raising AFDC payments early in life has a lasting impact on labor supply later in life.[16] While the

[14]Variables with a prefix of MISS are dummy variables indicating whether or not the variable following the prefix is missing. Therefore, the sample mean is the proportion of missing observations for the named variable.

[15]Following a suggestion by Derek Neal in the discussion at the Conference, we excluded AFQT from our logistic regressions on high school completion, because of potential endogeneity problems arising from the joint determination of AFQT scores and high school completion rates.

[16]We also regressed hours worked by women on the mean number of hours that men worked in the state, plus the other variables listed for the cross-sectional estimator. This

Table 4:
Background Characteristics of All Women and Poor Women
(weighted to the 1978 U.S. population)

	All Women		Women Family Income < $10,000	
Variable	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
AGE 14	0.089	0.283	0.079	0.244
AGE 15	0.128	0.332	0.100	0.273
AGE 16	0.125	0.329	0.099	0.271
AGE 17	0.128	0.332	0.110	0.284
AGE 18	0.125	0.330	0.144	0.319
AGE 19	0.131	0.335	0.145	0.320
AGE 20	0.118	0.321	0.138	0.313
AGE 21	0.127	0.331	0.152	0.326
BLACK	0.147	0.352	0.225	0.379
HISPANIC	0.066	0.247	0.089	0.259
RIMG	0.045	0.207	0.050	0.199
PARIMG	0.047	0.210	0.056	0.209
MISS-PARIMG	0.018	0.132	0.030	0.156
INTACT	0.746	0.433	0.654	0.431
FEMHEAD	0.129	0.333	0.200	0.363
MOMED	11.015	3.657	10.396	3.652
MISS-MOMED	0.052	0.221	0.0719	0.234
DADED	10.576	4.935	9.4542	4.941
MISS-DADED	0.111	0.312	0.1718	0.342
AFQT89	0.261	0.932	0.0588	0.837
MISS-AFQT89	0.056	0.230	0.0721	0.234
RAFDC	426.80	176.39	417.47	170.25
PCTMANU	0.200	0.064	0.198	0.057
FAMWEL78	0.112	0.315	0.138	0.313
OBSERVATIONS	4885		2369	

Table 5:
Logit for High School Completion: All Women

Variable	Parameter Estimates (Standard Errors in Parenthesis)		
RAFDC	0.0000	-0.0006	-0.00035
	(0.0003)	(0.0015)	(0.000279)
INTERCEPT	-2.3794	-4.5483	-4.0846
	(0.3958)	(1.7507)	(0.5516)
BLACK	0.1402	0.0877	0.1703
	(0.1204)	(1.271)	(0.1207)
HISPANIC	0.0925	0.1172	0.1120
	(0.1732)	(0.1838)	(0.1731)
RIMG	-0.0171	-0.0163	-0.0076
	(0.2732)	(0.2789)	(0.2735)
PARIMG	0.5215	0.5284	0.5520
	(0.2826)	(0.2872)	(0.2830)
INTACT	0.9709	0.9777	0.9574
	(0.1115)	(0.1143)	(0.1118)
FEMHEAD	0.3342	0.3557	0.3305
	(0.1405)	(0.1430)	(0.1409)
MOMED	0.1949	0.1951	0.1923
	(0.0199)	(0.0202)	(0.0198)
DADED	0.0960	0.0951	0.0945
	(0.0164)	(0.0166)	(0.0164)
PCTMANU	2.6177	11.3233	1.3716
	(0.6938)	(6.4693)	(0.7528)
FAMWEL78	-0.2783	-0.2600	0.2612
	(0.1225)	(0.1244)	(0.1229)
COHORT EFFECT	YES	YES	YES
OPPOSITE SEX CONTROLS	NO	NO	YES
STATE FIXED EFFECT	NO	YES	NO

Table 6:
Logit for High School Completion:
Women with Family Income Less Than $10,000 in 1978

Variable	**Parameter Estimates (Standard Errors in Parenthesis)**		
RAFDC	-0.00006	-0.00102	-0.00014
	(0.0035)	(0.00218)	(0.000367)
INTERCEPT	-1.7912	-4.1582	-2.1702
	(0.5596)	(2.6584)	(0.6686)
BLACK	0.1486	0.1148	0.1656
	(0.1506)	(0.1602)	(0.1516)
HISPANIC	-0.0466	-0.0959	-0.0270
	(0.2237)	(0.2415)	(0.2245)
RIMG	0.0840	0.0143	0.0672
	(0.3711)	(0.3905)	(0.3716)
PARIMG	0.4921	0.5207	0.5179
	(0.3565)	(0.3686)	(0.3569)
INTACT	0.9481	0.9628	0.9492
	(0.1516)	(0.1571)	(0.1517)
FEMHEAD	0.4979	0.5393	0.4923
	(0.1800)	(0.1856)	(0.1801)
MOMED	0.1414	0.1444	0.1402
	(0.0260)	(0.0269)	(0.0260)
DADED	0.1286	0.1261	0.1295
	(0.0228)	(0.0234)	(0.0228)
PCTMANU	1.1750	11.1755	0.9899
	(0.9748)	(9.8052)	(0.9932)
FAMWEL78	-0.0789	-0.0753	-0.0773
	(0.1665)	(0.1703)	(0.1663)
COHORT EFFECT	YES	YES	YES
OPPOSITE SEX CONTROLS	NO	NO	YES
STATE FIXED EFFECT	NO	YES	NO

cross-sectional and fixed-effect models suggest that raising AFDC payments early in life lowers labor supply in the late twenties the magnitude of the effect is quite different.[17] There is a large literature that demonstrates that higher AFDC payments increase the take-up rate in AFDC.[18] Furthermore, almost no women on AFDC work.[19] One plausible explanation is that reducing AFDC payments would increase labor-supply market experience in early years and induce greater labor supply in later years. On this view, taking AFDC early in life has a lasting effect after the support has ended, but conditioning on previous labor supply eliminates the correlation between past AFDC participation on current labor supply.

Table 8 presents the effect of AFDC generosity at age 17 from the cross-sectional and state fixed-effect estimator controlling for the background variable above and hours worked at ages 24 through 26.[20] While Table 7 suggests that when we do not control for previous hours worked, more generous AFDC payments lower labor supply, as Table 8 shows, the evidence is much weaker once we control for work experience. The cross-sectional estimator suggests that there is no direct effect of past AFDC generosity on labor supply. With conditioning upon past work experience, the state fixed-effects estimator now reverses the sign of the effect of past AFDC on labor supply at age 27. That is, controlling for hours worked at ages 24 through 26, women who lived in states at age 17 with higher AFDC benefits work more at age 27. While this seems counterintuitive, Hotz, McElroy, and Sanders (1995) found a comparable result for the NLSY, that teenage childbearing reduces labor supply early in life but significantly raises labor supply by age 27.

4 A simple model of wage growth, work, and welfare participation

4.1 *Some intuition*

We now present a simple dynamic model of work and welfare participation which forms the basis for our structural estimation. The essential features of the model are that wage growth comes from working, and that women

opposite sex estimator yields similar results to the cross-sectional estimator, perhaps because the extra regressor was found statistically insignificant in the unreported regression.

[17]The mean number of hours worked at age 27 is 550 and the mean real AFDC payment is \$430. The fixed-effect estimate suggests a 10-percent change in AFDC generosity leads to an implausibly large 73-hour, or 13-percent, decline in labor supply in the population.

[18]See O'Neill, Wolf, Bassi, and Hannan (1984), and Bane and Ellwood (1985).

[19]Prior to 1983 a women could work while on the AFDC program, but her AFDC grant was reduced by \$0.66 for every dollar she earned. After 1983 all earned income (with the exception of a disregard of \$30.00) was subtracted from her AFDC grant.

[20]We also investigated controlling for hours worked at each age between ages 18 and 26 and found similar results.

Table 7:
The Effect of AFDC Generosity at Age 17 on Hours Worked at Age 27

Variable	**Parameter Estimates (Standard Errors in Parenthesis)**		
RAFDC	-0.22	-1.70	-0.22
	(0.12)	(0.64)	(0.12)
INTERCEPT	364.60	788.06	355.25
	(177.76)	(737.88)	(194.66)
BLACK	-10.07	-57.13	-9.63
	(59.94)	(62.30)	(60.19)
HISPANIC	-57.81	-47.62	-57.89
	(86.67)	(90.88)	(86.66)
RIMG	18.30	-32.37	-18.00
	(121.19)	(121.69)	(120.81)
PARIMG	123.70	107.73	123.57
	(121.99)	(122.84)	(122.10)
INTACT	-124.83	118.96	-124.45
	(58.77)	(59.21)	(58.87)
FEMHEAD	-179.63	-171.63	-179.40
	(74.85)	(75.14)	(74.87)
MOMED	19.40	19.52	19.36
	(9.17)	(9.25)	(9.18)
DADED	4.87	4.65	4.80
	(7.07)	(7.12)	(7.07)
AFQT89	15.10	13.39	15.22
	(24.75)	(25.07)	(24.79)
PCTMANU	14.06	950.32	22.31
	(299.15)	(319.76)	(305.62)
KID 1	-211.28	-214.77	-211.16
	(55.42)	(55.78)	(55.44)
KID 2	-167.35	-162.04	-167.38
	(29.37)	(29.49)	(29.38)
KID 3	-62.18	-54.04	-62.19
	(31.97)	(32.11)	(31.97)
KID 4	-22.01	-32.65	-21.92
	(119.62)	(119.60)	(119.78)
FAMWEL78	-80.58	-76.57	-80.63
	(60.40)	(60.67)	(60.40)
COHORT EFFECT	YES	YES	YES
OPPOSITE SEX CONTROLS	NO	NO	YES
STATE FIXED EFFECT	NO	YES	NO

Table 8:
The Effect of AFDC Generosity at Age 17 on Hours Worked at Age 27 after Conditioning on Past Labor Supply

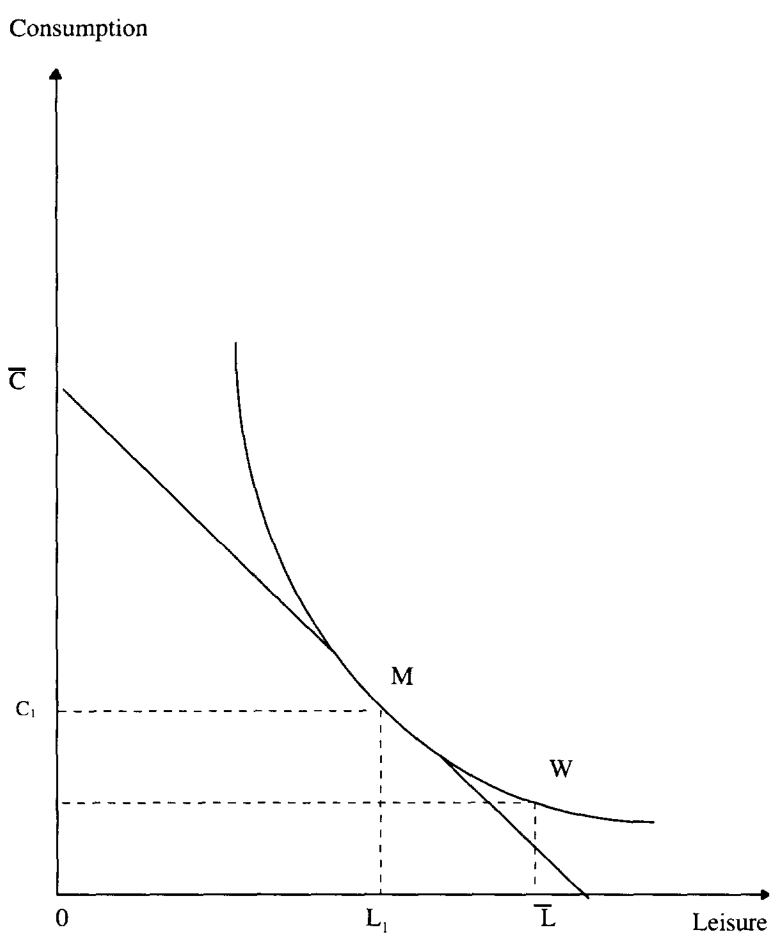
Variable	**Parameter Estimates (Standard Errors in Parenthesis)**	
RAFDC	-0.11	1.20
	(0.11)	(0.56)
HOURS AT AGE 24	0.04	0.03
	(0.01)	(0.01)
HOURS AT AGE 25	0.12	0.12
	(0.01)	(0.01)
HOURS AT AGE 26	0.22	0.22
	(0.01)	(0.01)
COHORT EFFECT	YES	YES
OPPOSITE SEX CONTROLS	NO	NO
STATE FIXED EFFECT	NO	YES

anticipate this in making current work and welfare choices. The effects of these features can be illustrated diagrammatically with reference to a one-period model. In Figure 2 leisure is measured along the horizontal axis, and consumption expenditure, net of inheritances and bequests, is measured on the vertical axis. In the absence of assistance, the budget set is the triangle with coordinates 0, $\bar{L}$, and $\bar{C}$, and in a one-period model the woman chooses M, which yields leisure and consumption of L_1 and C_1, respectively. Suppose the government introduces a transfer scheme (paid for by other taxpayers) which guarantees a minimal income with a dollar-for-dollar benefit reduction with earned income. Then the maximal benefits must be at least W to induce the woman to leave her employment and take public assistance. Notice that in a repetition of a one-period model like this one, where all income must be spent when it is earned, women on welfare never leave it, and women who do not take welfare at the beginning of their lives never choose it. The reason for making such obvious points from such a simple model is to emphasize that dynamic features need not be an important factor in explaining why long spells of welfare occur, and why early career decisions are so revealing about the future.

Suppose that the benefits exceed W but rather than go on public assistance, the woman whose one-period indifference curve is illustrated in Figure

FIGURE 2

A Static Model of Welfare Participation



2, continues to work. The explanation most commonly given in the literature is that there is a stigma attached to going on welfare which effectively reduces its net benefits below the payout. In this case point M accurately reflects the utility from working, but W overstates the utility from going on welfare. Alternatively, the wages earned in the current period may not fully represent the benefits from working. In this case the point M understates the utility from working, but W truly reflects the utility from going on welfare. In either case, with maximal benefits set at W, the woman would strictly prefer working to welfare.

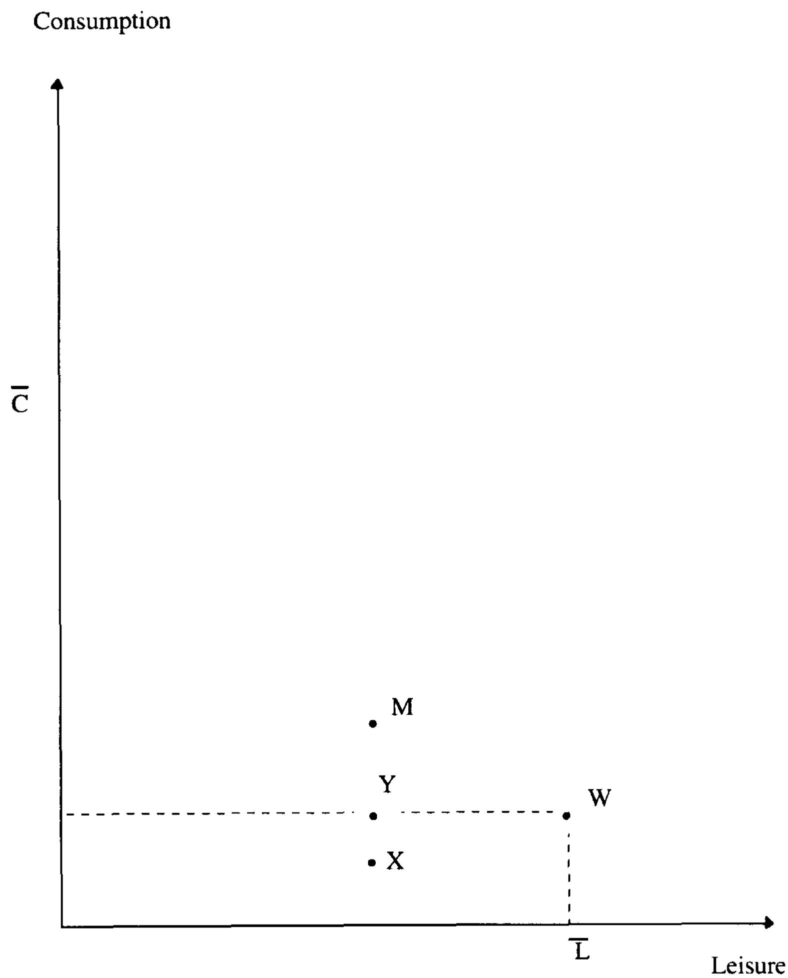
Of course utility functions are not part of the data base, and it might not seem possible to identify the role of dynamics from stigma. There are essentially three reasons why a person might choose M over W in Figure 3. First the one-period indifference curve through W might pass to the left of M. Second, the one-period indifference curve through W might pass to the right of M, but the person may face a sufficiently high stigma cost to deter her from collecting welfare; third there might be benefits from working that accrue to the person in the future, inducing her to stay on the job even though her utility is lower in the current period. But not all our data consist of women who locate at W or at points like M which lie strictly to its northwest. They also choose points such as X and Y. If a person chooses X, foregoing both leisure and consumption, we would conclude that stigma has a real effect (because if there was no stigma attached to receiving welfare they could increase their utility by increasing their consumption without adjusting their labor supply), whereas if a person chooses Y, by simultaneously working but enrolling in welfare as well, then we should infer that the benefits of working exceed the benefits of leisure foregone by more than the current wage rate. (If working did not confer human capital, a woman would prefer W to Y because W involves no work.) These observations prompted us to estimate a structural model which incorporates stigma and also includes dynamic considerations.

4.2 *The human capital production process*

If human capital production is an explanation for welfare dependence and for nonparticipation of welfare-eligible women, then wages must increase with greater labor-market experience to potentially explain welfare nonparticipation. Similarly, wages should decline with absence from work, if recent work experience raises wages more than work experience gained in the distant past. These hypotheses are consistent with a notion of human capital that accumulates as a by-product of working, but decays with nonuse. If true, then when a women stops working, her stock of effective human capital falls as her skills become rusty through neglect. In a steady state, the stock of effective human capital is balanced at the point where hours spent at work

FIGURE 3

Revealed Preferences for Welfare Participation



are just sufficient to maintain the current skill level. Therefore, the higher the labor supply, the greater the steady-state level of human capital.

Denote the wage of the i^{th} woman at time t by w_{it}. After experimenting with several functional forms, we decided to represent the learning-by-doing process by the equation:

$$ln(w_{it}) = ln(\delta_i) + \sum_{m=1}^{5} \lambda_m(Exp_{imt}) + \varphi g(b_{it}) + \eta_{it} \tag{1}$$

where $g(b_{it})$ captures time-varying factors that influence wages, including total labor-market experience, age, education, and local labor-market conditions, while Exp_{imt} is a measure of how total labor-market experience is distributed across past periods. In theory, Exp_{imt} might be any function of the history accumulated up until the current time period, but a detailed discussion of how we defined this variable for our empirical work is deferred until Section 5.

4.3 *Choices and preferences*

The teenager is assumed to choose, month by month, whether to work or whether to participate in welfare. Let $P_{it} \in \{0,1\}$ denote whether the i^{it} woman takes welfare ($P_{it} = 1$) or not ($P_{it} = 0$) at time t. Also let $L_{it} \in \{0,1\}$ denote whether the i^{it} woman works ($L_{it} = 1$) or not ($L_{it} = 0$). The four permutations can be represented by indicator variables d_{lit} through d_{4it}, defined as:

$$d_{lit} = \begin{cases} 1 \text{ if } P_{it} = 0 & \text{and } L_{it} = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$d_{2it} = \begin{cases} 1 \text{ if } P_{it} = 1 & \text{and } L_{it} = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$d_{3it} = \begin{cases} 1 \text{ if } P_{it} = 0 & \text{and } L_{it} = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$d_{4it} = \begin{cases} 1 \text{ if } P_{it} = 1 & \text{and } L_{it} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

These indicator variables are both exhaustive and exclusive, in the sense that for each i and t:

$$\sum_{j=1}^{4} d_{jit} = 1 \tag{3}$$

Consolidating the notation with $d_{it} \equiv (d_{lit}, d_{2it}, d_{3it}, d_{4it})$, we assume that a woman sequentially chooses $\{d_{is}\}_{s=0}^{T}$ to maximize her lifetime expected utility, here defined as:

$$E_0\{\sum_{s=0}^{T}\beta^s\sum_{j=1}^{4}d_{jis}[u_j^*(c_{jis},h_{is})+\epsilon_{jis}]\} \tag{4}$$

where β is the discount rate of utility, E_t is the expectation operator at time t, c_{js} the consumption associated with choice $j = 1, ..., 4$ for each period $s = t, ..., T$, the woman's history (including her initial endowment, her past labor-force and welfare participation choices, as well as births and marriages) up to time s is h_{is}, and ϵ_{jis} is an unobserved, choice specific variable. Stigma and the disutility of work directly affect a woman's current utility function. Welfare addiction and habit formation of a work ethic enter utility through her history, h_{is}. We assume the woman observes ϵ_{jis} for all four alternatives in the current period before making her choice for that period, but that she does not observe future state-specific utility shocks before making her current choice. However, she does know $G(\epsilon_{1js}, \epsilon_{2js}, \epsilon_{3js}, \epsilon_{4js})$, the probability distribution from which the vector $(\epsilon_{1js}, \epsilon_{2js}, \epsilon_{3js}, \epsilon_{4js})$ is drawn.

There are several ways one could model the financial and labor-market opportunities available to young women. Altug and Miller (1995) assumed there are no impediments to achieving an efficient allocation of resources and that hours supplied are a continuous choice variable. Here we assume that there are no opportunities to spread wealth across periods, and that the choice over hours worked is very restricted, assumptions which are much closer in spirit to Eckstein and Wolpin (1989). Thus, the woman's budget constraint in each period is described by:

$$c_{jit} = \begin{cases} 0 & \text{if} \quad d_{lit} = 1 \\ 160w_t & \text{if} \quad d_{2it} = 1 \\ G & \text{if} \quad d_{3it} = 1 \\ G + \tau 40 w_t & \text{if} \quad d_{4it} = 1 \end{cases} \tag{5}$$

The budget constraint assumes that if a woman is not on welfare and works, she works full-time (160 hours per month) and earns $160w_t$. Of course, if she is neither on welfare nor works she has no earnings at all. A woman on welfare receives a cash grant, G, but only keeps a fraction, τ, of her earnings. The budget constraint assumes that women on welfare who work, work only 10 hours a week (40 a month). Abstracting from the hours' decision greatly simplifies estimation and seems fairly innocuous given that most women who work and are also on welfare work close to 10 hours per week.

Notice that a woman's past choices, h_{it}, affects the budget constraint both through wages and through the propensity to work and use welfare. We assume that wages follow Equation (1). We suspect that past work experience raises current wages, with more recent work experience raising current wages more than more distant work experience, that is, $\lambda_1 > \lambda_2 > ... > \lambda_5$. Since past choices affect current utility, past choices will also affect the propensity to work and to use welfare. This in turn affects current consumption.

4.4 *Optimal choices and the conditional choice probabilities*

The teenager must choose d_{it} to maximize the objective function described by Equation (4) subject to the budget constraint described by Equation (5). Let d^0_{it} denote the optimal decision in period t. Bellman's (1957) principle implies that a choice k is optimal, that is, $d^0_{kt} = 1$, if:

$$k = arg \max_{j \in \{1,2,3,4\}} [u^*_{jit} + \epsilon_{jit} + v_j(h_{it})] \tag{6}$$

where $u^*_{jis} \equiv u^*_j(c_{jis}, h_{is})$ and the conditional value function for j, denoted $v_j(h_{it})$, is defined as:

$$v_j(h_{it}) \equiv E_t\{ \sum_{s=t+1}^{T} \beta^{s-t} \sum_{j=1}^{4} d^0_{jis}(u^*_{jis} + \epsilon_{jis}) | h_{it}, d_{jit} = 1\} \tag{7}$$

Equation (6) states that k is the optimal choice if alternative k yields the largest sum of current and future utility. Equation (7) states that the future utility for each choice j is simply the discounted future stream of period-specific utility that accrues to choice j when all future choices are made optimally.

The probability that choice k is optimal, given a woman's history of past choices and outcomes, h_{it}, is given by:

$$p_k(h_{it}) = P_r\{k = arg \max_{j \in \{1,2,3,4\}} [u^*_{jit} + \epsilon_{jit} + v_j(h_{it})] | h_{it}\} \tag{8}$$

That is, conditional on h_{it}, the probability of observing the i^{th} woman make a particular work/welfare choice, k, is just the probability that the other choices do not yield as much current and future utility as k. Since the conditional value functions only depend h_{it}, on a vector observed by the econometrician, it follows that a maximum likelihood estimator could be formed in the usual manner, if the functional form of the conditional value functions were known, or could be easily computed numerically. Unfortunately this is rarely the case because the decision tree underlying the agent's problem quickly becomes intractable for all but the simplest frameworks.

Rather than solve the optimal decision rule directly for each set of candidate values in a likelihood function, we exploit results in Hotz and Miller (1993) which establish a one-to-one mapping between the conditional choice probabilities and the conditional valuation functions. This relationship is key to the simulation estimator developed in Hotz, Miller, Sanders, and Smith

(1994) applied here. This approach requires us to first compute the conditional choice probabilities, and the transition probabilities, before inferring the structural parameters. Our results on structural estimation are laid out in three parts. Our findings on the wage equation are reported first, in the next section. Section 6 describes some of the implications that come from our estimates of the conditional choice and transition probabilities. Finally the structural parameter estimates are reported in Section 7.

5 Endogenous wage growth

To model the experience variable in the wage equation, we decomposed the Exp_{imt} variables into labor-market experience accumulated over five time periods $m = 1, ..., 5$. These categories represent the number of months worked in the previous four months ($m = 1$), the number worked in the previous five to twelve months ($m = 2$), the number of months worked in the previous one to two years ($m = 3$), between two to three years ($m = 4$), and between three to five years ($m = 5$).[21] Thus, the λ coefficients in Equation (1) represent the relative value of work experience accumulated at different times in the past. We are specifically interested in investigating three hypotheses: whether the coefficient on total labor-market experience is positive, whether the λ coefficients are positive, and whether they decline monotonically, that is, whether $\lambda_1 > \lambda_2 > ... > \lambda_5$. Evidence on the first hypothesis would help answer whether negative duration dependence on welfare can be attributed to the process of human capital production, while the second two hypotheses test whether recent work experience raises wages more than work experience in the distant past (or that human capital decays over time through disuse).

The model incorporates a person-specific component to wages, δ_i. We assume that selection in labor-force participation occurs only on this person-specific component, uncorrelated measurement error accounting for the remaining unobserved heterogeneity. This assumption implies that properly controlling for δ_i would eliminate the selection problem. Accordingly, we treated δ_i as a fixed effect, estimating Equation (1) in first differences to remove this systematic source of unobserved heterogeneity and potential for selection bias. The resulting equation was estimated on a 25-percent sample of the 275,265 months in which women sampled from the NLSY worked.

Table 9 presents the coefficients on the 5 experience categories, the coefficient on the return to total labor-market experience, and their estimated

[21]Since $g(b_{it})$ controls for total labor-market experience, the final category, the number of months worked more than five years ago is omitted from Exp_{imt} in the wage regression. In addition, share-weighted changes in annual sectoral employment growth within the county where the woman lives were entered into the wage equation as factors affecting $g(b_{it})$.

errors. An additional month of labor-market experience raises wages by about 0.3 percent. However, if a woman gained this experience within less than four months ago, the return is closer to 1.7 percent. In contrast, if this experience was obtained three to five years ago, wages would on average increase well less than 0.1 percent (.003207 - .002811). For the most part recent work experience raises wages more than work experience in the distant past, that is, $\lambda_m > \lambda_{m+1}$ for all $m = 1, ..., 4$, a finding confirmed by Altug and Miller (1995) in their empirical study using the PSID. We take this as evidence that human capital decays over time.

Table 9:
Coefficients on Labor-Market Experience in Wage Equation
(with controls for education, age and local labor markets)

Variable	**Parameter Estimates (Standard Errors in Parenthesis)**
TOTAL EXPERIENCE	0.003307
	(0.005596)
EXPERIENCE 1 - 4	0.014159
	(0.001467)
EXPERIENCE 5 - 12	0.007828
	(0.000693)
EXPERIENCE 13 - 24	0.003420
	(0.000504)
EXPERIENCE 24 - 36	-0.001445
	(0.000475)
EXPERIENCE 37 - 60	-0.002811
	(0.000389)
OTHER BACKGROUND VARIABLES	YES

However, the negative estimates obtained for λ_4, λ_5, and the coefficient on total experience show that although experience gained from working two to five years has a positive effect on wages (which is deduced by noting that the effect of total experience is of opposite sign and greater in magnitude), work experience gained from longer than five years ago raises wages even more. We are skeptical about this literal interpretation of our findings; another interpretation is that unobserved heterogeneity is not fully captured by the δ_i fixed effects, and that women with work experience dating back beyond five years are drawn from a different, more career-oriented population then the others. We conjecture that if the δ coefficients differ across individuals as well as across lagged experience, misspecifying the model by ignoring these differences might lead to the pattern of estimates we have found.

Figure 4 simulates two wage paths. The first is a typical wage path for a career woman, and the second for a welfare recipient.[22] The top panel (ContWork) represents predicted wages for a woman who worked continuously from age 16 to 25. The bottom panel (StopWork) represents the predicted wages for a woman who worked continuously from age 16 to 19, but then stopped working. (Age 19 is the age at which women in the NLSY are most at risk to first enter welfare.) The simulation in Figure 9 shows that leaving the labor force lowers a woman's wages not only relative to what they would have been had she worked, but also relative to the wage she might have received when she stopped working. After a 19-year-old woman stops working, she can expect her wages to fall from \$4.27 per hour to \$3.60 per hour over the next two years, about a 16-percent decline.

6 Findings from nonparametric estimation

When the i^{th} woman enters period t with a history, h_{it}, the vector ϵ_{it} is revealed to her, and she makes some optimal work-welfare choice, d^0_{it}. Having made her work-welfare choice, one of several outcomes might occur by the end of time t.[23] The woman could have worked or not, be married or not, be pregnant or not, and if she was pregnant in the previous period she could lose the child or have a birth by the end of t. Define this outcome vector as b_t, where a typical element b_{jt} represents a particular labor participation choice, marriage outcome, and fertility/pregnancy outcome in period t. The woman does not know which marriage or pregnancy outcome will occur, but she does know the probability that each outcome occurs conditional on her history and her choice, $F(b_{it}|h_{it}, d_{it})$, the cumulative probability distribution of outcomes. Since a woman can be married or not, and can be pregnant or not, the vector b_{it} represents the four marriage-pregnancy states that might occur. Our assumption about the time independence of the unobserved variables implies $h_{i,t+1} \equiv (h_{it}, b_{it})$.

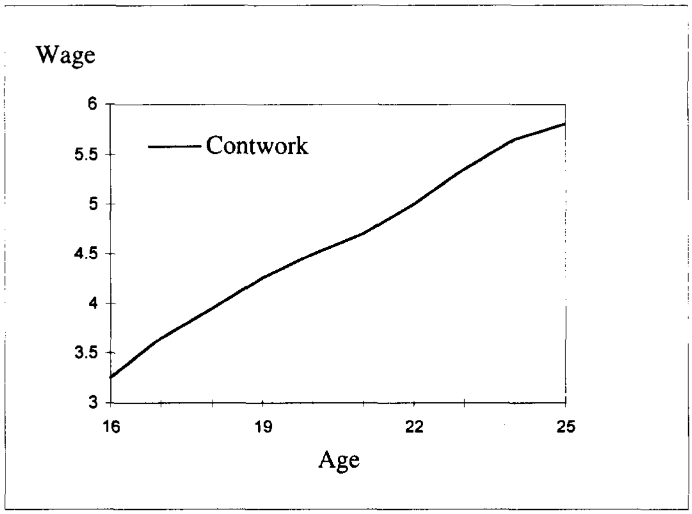
Although the data only record h_{it} ad d^0_{it}, but not ϵ_{it}, the assumption that ϵ_{it} is independent over time implies the choice and transition probabilities that an econometrician estimates to characterize the woman's future are those the woman herself uses when planning her own future. Furthermore, Hotz and Miller (1993) show that the expected value of the disturbance

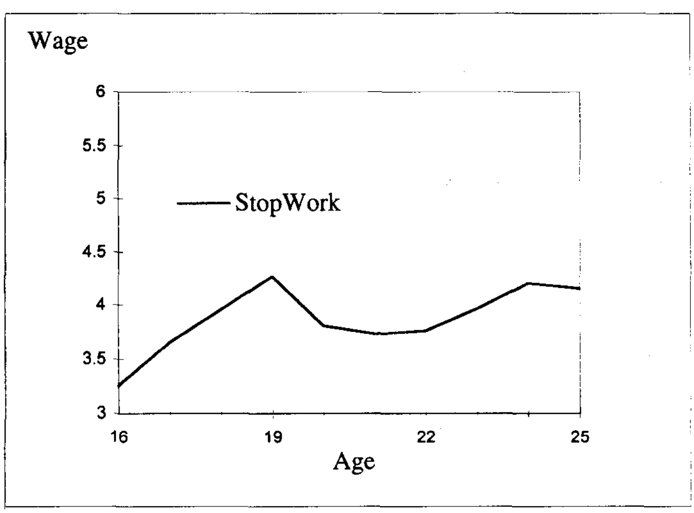
[22]The simulation is necessary since age and the six-part spline on total experience will also change wages as labor-market experience is accumulated. If, for example, a positive age effect is large enough to allow wages to rise even in the presence of decaying human capital, then the wage process could not explain welfare dependence. In this simulation we assumed the fixed effect is the estimated sample average, and that the average labor-market conditions found over the survey prevailed.

[23]The timing convention is that in a month t, the outcome that ended period $t-1$ and the resulting choice in t are observed. The outcome that occurs at the end of t is recorded at the beginning of period $t+1$.

FIGURE 4

Predicted Wages in $1982 as a Function of Age





associated with the optimal choice can be written as a function of the conditional choice probabilities. In other words, defining the vector of conditional choice probabilities by $p(h_{it}) \equiv (p_1(h_{it}), ..., p_3(h_{it}))$, there exists a mapping from the conditional choice probabilities, denoted $w_j[p(h_{it})]$ for each action j, such that:

$$w_j[p(h_{it})] = E[\epsilon_{jit}|d_{jit} = 1, h_{it}]. \tag{9}$$

Consequently we can write the conditional value functions as:

$$\begin{aligned} v_j(h_{it}) &\equiv E_t\{\sum_{s=t+1}^{T} \beta^{s-t} \sum_{j=1}^{4} d^0_{jis}(u^*_{jis} + \epsilon_{jis})|h_{it}, d_{jit} = 1\} \\ &= E_t\{\sum_{s=t+1}^{T} \beta^{s-t} \sum_{j=1}^{4} p_j(h_{is})(u^*_{jis} + w_j[p(h_{is}])|h_{it}, d_{jit} = 1\} \end{aligned} \tag{10}$$

where the expectation in the bottom line of (10) is taken by integrating over the transition probabilities $F(b_{it}|h_{it}, d_{it})$.

It follows from these remarks, and Equation (10), that knowledge of $p(h_{it})$ and $F(b_{it}|h_{it}, d_{it})$ essentially characterizes the solution to the agent's problem. For this reason we investigated the stochastic behavior generated by these probability distributions before imposing any parametric assumptions on preferences. We estimated $F(b_{it}, |h_{it}, d_{it})$ and $p_j(h_{it})$ using nonparametric regression techniques; a kernel estimator was used to calculate the distance between a particular woman and a sample of matching women. Having obtained estimators for the conditional choice and transition probabilities, we then simulated paths that women living in this economy would take.[24] Averaging over these future simulated paths allows one to display in a nonparametric way the future propensity to work and use welfare conditional on current choices. The results from this exercise are plotted in Figures 5 and 6.

Figure 5 presents the average simulated path of monthly labor-force participation a woman could take in the future, conditioning upon a set of initial work-welfare choices. For example, the circles represent the typical future work path if a woman initially chooses to work ($L = 1$) and stay off welfare ($P = 0$), while the triangles represent the typical future work path if a woman does not work ($L = 0$) and enters AFDC ($P = 1$) in the current period. Compared to a woman who works and does not enter AFDC, a woman who does not work and enters AFDC remains less likely to work many months later. While the predicted gap in the propensity to work narrows over time, a 10-point difference remains after 5 years. Interestingly, women who neither work nor enter welfare in the current period (straight line) are more likely

[24]The details of this procedure are found in Hotz, Miller, Sanders, and Smith (1994).

to work in the future than women who work in the current period but also enter AFDC in the current period (squares).

Figure 6 shows the likely reason. It displays the future propensity to enter AFDC for each of the 4 current work-welfare alternatives. The figure shows that current welfare use is positively associated with future welfare use, and that once the decision has been taken to accept welfare, the choice about whether or not to work hardly affects the future propensity to use the AFDC program. Together, Figures 5 and 6 show that women who enter welfare and work soon stop working altogether, leaving them solely dependent on welfare in the future. Thus, 75 percent of the women who enter welfare working remain on welfare for at least a year, but only 25 percent of them are still working after a year has elapsed.

7 Parametric estimates of preferences

This section reports the results from estimating a parameterization of the dynamic model of welfare participation we presented in Section 4, and testing for the importance of stigma and habit formation. The estimates presented below assume the observed part of the monthly utility function takes the form:

$$\begin{aligned} u_j^*(c_{jis}, h_{is}) &= \gamma_{1j} c_{jis} - \gamma_{2j} c_{jis}^2 + \psi X_{1is}(d_{2is} + d_{4is}) + \alpha X_{2is}(d_{3is} + d_{4is}) \\ &= \gamma_{1j} Y_{jis} - \gamma_{2j} Y_{jis}^2 + \psi X_{1is} P_{jis} + \alpha X_{2is} L_{jis} \qquad (11) \end{aligned}$$

where X_{1is} and X_{2is} are vectors of background characteristics that are determined by the history h_{is}, α and ψ are conformable parameter vectors to be estimated, while γ_{1j} and γ_{2j} are choice-specific parameters that together characterize the marginal utility of consumption for each $j \in \{2, 3, 4\}$. Here utility from each choice is benchmarked against $d_{1is} = 1$ (that is, not receiving welfare, $P_{jis} = 0$, and not participating in the labor force, $L_{jis} = 0$), in which case $u_{1is}^* = 0$ (since nonlabor income not received from AFDC payments are ignored). The second line is derived from the first by noting from Section 4 that $L_{jis} = (d_{3is} + d_{4is})$ and $P_{jis} = (d_{2is} + d_{4is})$ for each $j \in \{1, 2, 3, 4\}$, while the assumption that there are no opportunities to save or borrow allows us to replace c_{jis} with Y_{jis}. Thus, αX_{2is} is a factor loading on L_{jis} which represents the marginal disutility of work, ψX_{1is} represents the disutility from participating in AFDC, while γ_1 and γ_2 together determine the marginal utility of current income (and also consumption).

The X_{1is} vector contains characteristics that affect the disutility of receiving AFDC, including an intercept to capture the base-line stigma effect (INTERCEPT), an indicator of whether the woman received welfare the previous month (AFDCLAG), the number of months the woman has been receiving AFDC since the current spell on welfare began (DURAFDC), a

FIGURE 5

Probability of Future Work given Current Work/Welfare (L,P) Choice at Time t

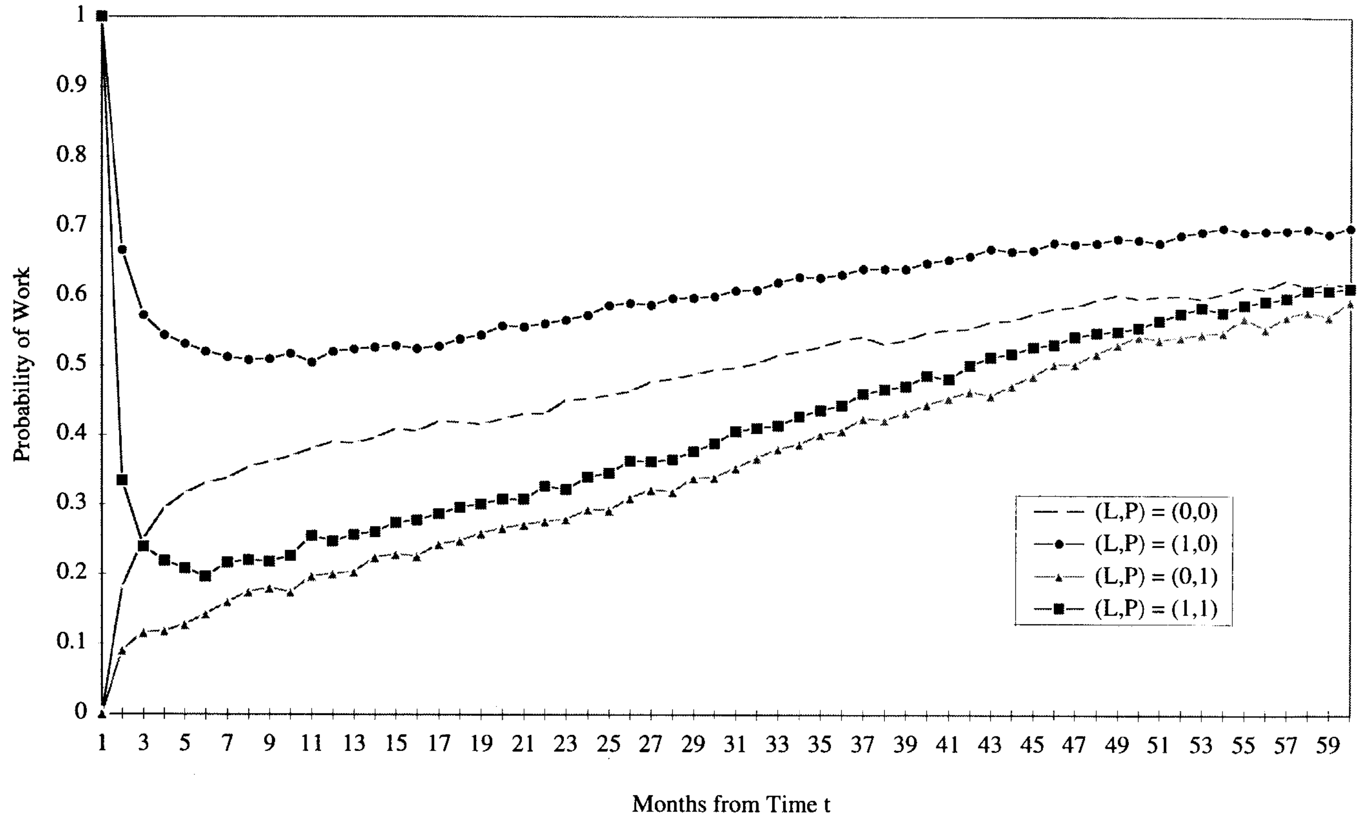
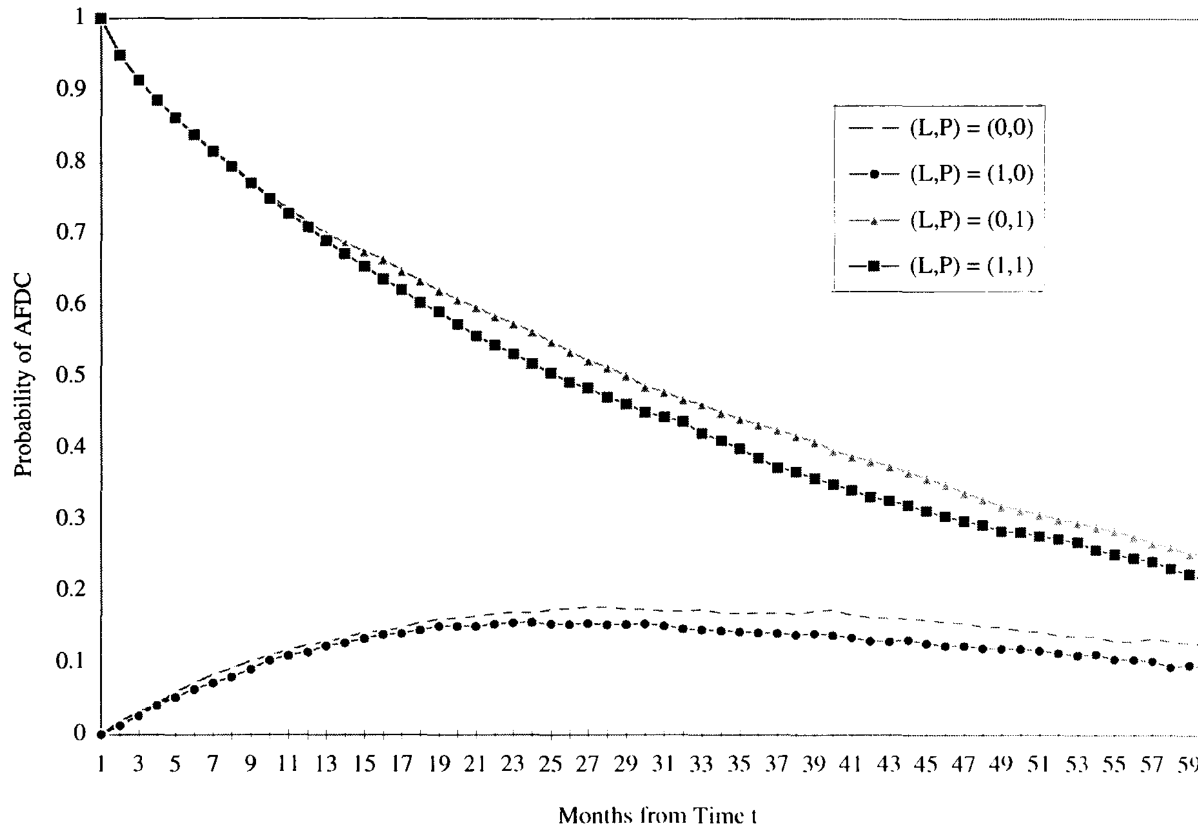


FIGURE 6

Probability of Future AFDC Given Current Work/Welfare (L, P) Choice



variable to control for the seam problem mentioned earlier (SPIKE), and whether the respondent's mother was on welfare in 1978 (FAMWEL78). In addition to INTERCEPT, AFDCLAG, DURAFDC, and SPIKE, the X_{2is} vector contains other characteristics that might affect the disutility of working, including an indicator variable for whether the respondent worked last period (WORKLAG), the number of months she worked in the previous twelve periods (SHORTEXP), the number she worked in the four years preceding that (LONGEXP), whether she is black (BLACK), whether she completed high school (HIGHSCHOOL), her age (AGE), the number of children at home (NUMCHILD), and the age of the youngest child at home (AGEYOUNGEST).

Our specification of preferences is completed by assuming here that the unobserved variable ϵ_{jis} is independently distributed across choices $j \in \{1,2,3,4\}$, sample observations $i \in \{1,...,I\}$, and time periods $s \in \{1,...,T\}$, as a type-1 extreme random variable. We remark that the expected value of the stochastic error term given i optimally selects alternative j at s, denoted $E[\epsilon_{ijs}|h_{is}, d_{ijs} = 1]$ exceeds $E[\epsilon_{ijs}|h_{is}]$, and in general differs from its unconditional expectation $E[\epsilon_{ijs}]$. For example, women are more likely to work and not participate in welfare in periods when the stochastic utility to working is large. Therefore, the expected value of stochastic utility when choices are made optimally is in general greater than the unconditional expectations. Hotz and Miller (1993) show this component can be calculated as a function of the estimated choice probabilities, and that when the stochastic error is distributed by an extreme value distribution, the expected stochastic utility associated with an optimal choice j is:

$$E[\epsilon_{ijs}|h_{is}, d_{ijs} = 1] = \zeta - ln(p_j(h_{is})) \tag{12}$$

where $\zeta = 0.5772166$ is Euler's constant and $p_j(h_{is})$ is the conditional probability that alternative j is optimal at time s given history h_{is}.

As we mentioned in Section 4, the estimation framework exploits an identity which relates the conditional choice probabilities to the conditional value functions. In particular, the orthogonality conditions which characterize our CCS estimator come from multiplying variables in the agent's information set at time s (which serve as instruments) with expressions that are based on the three-equation system:

$$\begin{aligned} 0 &= ln(p_{jis}/p_{0ia}) - u_j^*(c_{jis}, h_{is}) - u_0^*(c_{jis}, h_{is}) + v_j(h_{is}) - v_0(h_{is}) \\ &= ln(p_{jis}/p_{0is}) - \alpha X_{1is}L_{jis} + \psi X_{2is}P_{jis} + \gamma_{1j}Y_{jis} - \gamma_{2j}Y_{jis}^2 \\ &\quad +v_j(h_{is}) - v_0(h_{is}) \end{aligned} \tag{13}$$

for each work-welfare choice $j \in \{2,3,4\}$. Estimates of the conditional choice probabilities are obtained nonparametrically in the manner described in Section 6, and substituted into the left side of (13). The conditional value functions on the right side of the equation are replaced, by simulating, for each observation, a future path the agent might take (according to the non-parametrically estimated conditional choice and outcome probabilities), and evaluating the remaining lifetime utility the simulated path would generate using the utility function (11) and (12). The resulting estimator is consistent in the square root of sample size and asymptotically normal.[25]

Table 10 presents the results from estimating two models. Column 1 reports our findings when agents are assumed myopic, that is $\beta = 0$, while Column 2 presents estimates of a model where the monthly subjective discount factor, β, is set to 0.95. This still represents very heavy discounting of the future, because it implies an annual discount rate of over 50 percent. The models were estimated from data for the youngest women in the sample, aged 14 to 16 years old in 1979 who are on average between 18 and 19 years old while eligible for AFDC. As noted above in Table 1, a large fraction of very young women neither participate in AFDC nor work, most likely because family support is a substitute for both. Because family support is not explicitly modeled here, the estimated level of stigma might appear quite large, simply because many of these women are drawing on alternative sources of support.

Considering the first of the three panels, we see that the coefficient estimates on income and their underlying interpretations are quite sensitive to assumptions about the discount factor. For $\beta = 0$, at low income levels the marginal utility of income obtained from working is higher than the marginal utility of income obtained from welfare. The estimated linear coefficient on Y_{3is} (working and staying off welfare) exceeds that of Y_{4is} (working and receiving welfare), and both are greater than that on Y_{2is} (where welfare is the sole source of income). The underlying interpretation of this result is that the stigma attached to AFDC participation is not fixed but depends on the amount of assistance received. When we assume women are forward-looking, the results change markedly and they are supported by a very different interpretation. When $\beta = 0.95$, the estimated utility from income is highest when the woman receives welfare but does not work, and lowest if the woman receives income from both AFDC and market work. This finding is consistent with a household production function that exhibits complementarity between time spent at home and market goods in the household production function; women who receive welfare have more time and energy to prudently spend it than those who work, while women who do not specialize in one of these income-generating activities have even less opportunity to wisely husband their resources. The finding that the marginal utility of income is highest for

[25]See Hotz, Miller, Sanders, and Smith (1994) for a definition of the CCS estimator, it simplementation, proof of its large sample properties, and a Monte Carlo study of its finite sample properties.

Table 10:
CCS Parameter Estimates From Structural Model

	Variable	Parameter Estimates (Standard Errors in Parenthesis) $\beta = 0$	$\beta = 0.95$
Income	Y_{2is}	-0.000132	0.000258
		(0.00088)	(0.0000196)
	Y_{2is}^2	-0.000000877	-6.438E-08
		(0.0000012)	(5.42E-09)
	Y_{3is}	0.000044	0.000110
		(0.000116)	(0.0000143)
	Y_{3is}^2	5.95E-09	-1.884E-08
		(4.25E-08)	(3.35E-09)
	Y_{4is}	-0.016873	-0.000170
		(0.000734)	(0.0000181)
	Y_{4is}^2	0.000017877	4.931E-08
		(9.29E-07)	(4.82E-09)
Taste for AFDC	INTERCEPT	-5.779951	-2.187383
		(0.1344762)	(0.04213)
	AFDCLAG	12.206994	1.755269
		(0.066188)	(0.044415)
	DURAFDC	0.038718	0.015496
		(0.002463)	(0.000317)
	FAMWEL78	0.216446	0.038647
		(0.039935)	(0.00749)
	SPIKE	-1.421941	0.182435
		(0.099506)	(0.100239)

Table 10 continued:
CCS Parameter Estimates From Structural Model

	Variable	Parameter Estimates (Standard Errors in Parenthesis)	
		$\beta = 0$	$\beta = 0.95$
Taste for Work	INTERCEPT	-5.040363	-3.718608
		(0.314041)	(0.068584)
	AFDCLAG	1.342238	0.087084
		(0.053095)	(0.019973)
	DURAFDC	0.012855	0.004767
		(0.00243)	(0.00047)
	SPIKE	0.404938	0.244247
		(0.074711)	(0.062467)
	WORKLAG	3.159814	2.687868
		(0.060153)	(0.035929)
	SHORTEXP	0.180405	0.082563
		(0.007091)	(0.002288)
	LONGEXP	0.070564	0.003089
		(0.002573)	(0.000511)
	BLACK	0.135539	-0.006748
		(0.043864)	(0.008879)
	HIGHSCHOOL	0.363339	-0.004125
		(0.043935)	(0.008967)
	AGE	0.005375	0.000775
		(0.001739)	(0.00029)
	NUMCHILD	0.199408	0.007732
		(0.052201)	(0.010592)
	AGEYOUNGEST	-0.010731	-0.000662
		(0.001674)	(0.000321)

those women who derive all their income from welfare is also consistent with the empirical regularity that members of this group are more likely to receive support from other government welfare programs, such as food stamps and Medicaid, variables which our econometric specification omits.

Turning to the estimated coefficients for X_{1is}, which characterize preferences for welfare, the estimates in both Columns 1 and 2 display a strong and significant negative stigma effect.[26] In addition, there is evidence that the stigma effect lessens as women spend more time on the program. (Both AFDCLAG and DURAFDC are positive and significant.) From the results of the static model, one is tempted to conclude there is a strong stigma effect and substantial welfare addiction. Our results also confirm a consistent finding in the literature, that women whose mothers were once welfare recipients are more likely to be on welfare themselves.[27] This has led some researchers to speculate that there may be an intergenerational transition mechanism of welfare use, and that the stigma from welfare is smaller among women whose mothers were once themselves on welfare.

The second group of estimated coefficients apply to X_{2is} and display our estimates of the parameters characterizing leisure preferences. The negative and significant INTERCEPT for both models shows that, as theory would predict, the value of leisure is positive. It also seems that women build up a tolerance towards market work, or alternatively, that through experience working women acquire skills which help them to manage the competing demands made upon their time at home and at work. (WORKLAG, SHORTEXP, and LONGEXP are all positive and significant.) Notice that a large fraction of the disutility to work is overcome if a woman has worked in the prior month, and that the more a woman has worked in the past, the greater her taste for work becomes; more recent work experience is estimated to raise her taste for work more than more distant experience (that is, after controlling for its effect on wages). While these results hold for both the dynamic and the static models, the effects in the dynamic model are smaller in magnitude than the effects in the static model.

Many of the estimated coefficients on X_{1is} and X_{2is}, which respectively characterize the tastes for leisure and welfare enrollment, are an order of magnitude lower in the dynamic setup than the static one. The reason for this is straightforward. As Section 4 explained, in a dynamic model with nonseparable preferences, optimal behavior is determined by comparing the sum of current utility and the conditional value function associated with each

[26] The dollar value of the coefficients can be obtained by dividing the parameter estimates by the marginal utility of income. However, in the static model the marginal utility of income is measured quite imprecisely.

[27] See for example Antel (1991), Gottschalk (1989, 1990), McLanahan (1988), and Hill and Ponza (1984).

action. Then the action with the highest lifetime utility index is picked. If $\beta = 0$, the only way a variable affects its index is through current utility, whereas if the subjective discount factor is strictly positive, current utility captures only part of its effect, because there is an added effect to the index through the conditional value function. In particular, if past and current actions are complements, to induce the behavioral response observed in the data from a given variable change, the effect of a variable on current utility if $\beta = 0.95$ must typically be smaller than if $\beta = 0$. Recalling the evidence of duration dependence on and labor-supply welfare participation which we presented in the earlier parts of this paper, this seems to explain why many of the estimated parameters decline in magnitude when future utility is given added importance in decision-making by increasing β. Finally, this discussion is an empirical illustration of the critical role that the discount rate β plays in dynamic optimization. When β was estimated along with the taste parameters, an estimate greater than 1 was obtained. (The econometric criterion function is minimized at $\beta = 1.04$ and a t-test rejects the hypothesis that $\beta < 1$.) Athough this finding is indeed evidence that the best-fitting model exhibits forward-looking behavior, further empirical resarch is required to pin down the value of this key parameter.[28]

[28]The result that $\beta > 1$ has been found in several other empirical studies, including for example Hotz, Kydland, and Sedlacek (1988).

References

Altug, S. and Miller, R., (1995). The Effect of Work Experience on Female Wages and Labor Supply. *Review of Economic Studies*, forthcoming.

Antel, J., (1991) The Inter-Generational Transfer of Welfare Dependency: Some Statistical Evidence. Working Paper, University of Houston.

Bane, M. and Ellwood, D., (1985). The Impact of AFDC on Family Structure and Living Arrangements. *Research in Labor Economics*, **7**: 137-207.

Bellman, R., (1957). *Dynamic Programming.* Princeton University Press, Princeton.

Blank, R., (1989). Analyzing the Length of Welfare Spells. *Journal of Public Economics*, **39**: 245-273.

Eckstein, Z. and Wolpin, K., (1989). Dynamic Labor Force Participation of Married Women and Endogenous Work Experience. *Review of Economic Studies*, **56**: 375-390.

Fraker, T., Moffitt, R., and Wolf, D., (1985). Effective Tax Rates and Guarantees in the AFDC Program, 1967-82, *Journal of Human Resources*, **20**: 251-263.

Gottschalk, P., (1989). Correlation Between Mother's and Daughters Welfare Experiences. Working Paper.

Gottschalk, P., (1990). AFDC Participation Across Generations. *American Economic Review*, **80**: 367–371.

Hill, M. and Ponza, M., (1984). Does Welfare Dependency Beget Dependency? ISR Working Paper, University of Michigan.

Hotz, V., Kydland, F., and Sedlacek, G., (1988). Intertemporal Preferences and Labor Supply. *Econometrica*, **56**: 335–360.

Hotz, V. and Miller, R., (1993). Conditional Choice Probabilities and the Estimation of Dynamic Discrete Choice Models. *Review of Economic Studies*, **60**: 497–529.

Hotz, V., Miller, R., Sanders, S., and Smith, J., (1994). A Simulation Estimator for Dynamic Models of Discrete Choice. *Review of Economic Studies*, **61**: 265–289.

Hotz, V., McElroy, S., and Sanders, S., (1995). The Costs and Consequences of Teenage Childbearing for Mothers. Working Paper, NORC, 95-10.

Hoynes, H., (1995). Does Welfare Play Any Role in Female Headship Decisions? Working Paper, University of California Berkeley.

Jackson, C. and Klerman, J., (1994). Welfare, Abortion and Teenage Fertility. Working Paper, RAND.

McLanahan, S., (1988). Family Structure and Dependency: Early Transitions to Female Household Headship. *Demography*, **25**: 1-16.

Moffitt, R., (1983). An Economic Model of Welfare Stigma. *American Economic Review*, **73**: 1023–1035.

Moffitt, R., (1992). Incentive Effects of the U.S. Welfare System. *Journal of Economic Literature*, **30**: 1-61.

Moffitt, R., (1994). Welfare Effects on Female Headship with Area Effects. *Journal of Human Resources*, **29**: 621–636.

O'Neill, J., Wolf, D., Bassi, L., and Hannan, M., (1984). An Analysis of Time on Welfare. Final report to HHS, the Urban Institute.

Ruggles, P. and Michel, R., (1987). Participation Rates in the Aid to Families with Dependent Children Program: Trends for 1967 through 1984. Working Paper, The Urban Institute.